AFRL-IF-RS-TR-2007-131
Final Technical Report
May 2007

HARNESSING MULTIPLE REPRESENTATIONS FOR AUTONOMOUS FULL-SPECTRUM POLITICAL, MILITARY, ECONOMIC, SOCIAL, INFORMATION AND INFRASTRUCTURE (PMESII) REASONING

Rensselaer Polytechnic Institute

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

NOTICE AND SIGNATURE PAGE

AFRL-IF-RS-TR-2007-131 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/s/

PAUL BELLO
Work Unit Manager

/s/

JAMES W. CUSACK, Chief
Information Systems Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
MAY 2007	Final	Jan 06 – Jan 07

4. TITLE AND SUBTITLE
HARNESSING MULTIPLE REPRESENTATIONS FOR AUTONOMOUS FULL-SPECTRUM POLITICAL, MILITARY, ECONOMIC, SOCIAL, INFORMATION AND INFRASTRUCTURE (PMESII) REASONING

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA8750-06-2-0015

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Nicholas L. Cassimatis

5d. PROJECT NUMBER
230S

5e. TASK NUMBER
06

5f. WORK UNIT NUMBER
01

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Rensselaer Polytechnic Institute
110 8th St.
Troy NY 12180

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFRL/IFSB
525 Brooks Rd
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-IF-RS-TR-2007-131

12. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED. PA# 07-224

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Assessing the effects of various actions on particular varieties of battlespace entities has been impeded by the absence of a representative computational model of the governing dynamics between so-called Political, Military, Economic, Social, Information, and Infrastructure (PMESII) elements in the theater of operations. Associated with each of these categories are formalisms suitable for computational implementation. Currently, we are lacking a way to describe the nonlinear dependence of each category on the others. We have proposed that the formal systems used to describe each of these categories can be described within an abstracted, yet common representational framework and therefore heterogeneous reasoning across the PMESII spectrum may be within our grasp. The Polyscheme architecture (Cassimatis 2002) was utilized to perform this integration, and to demonstrate that non-trivial reasoning about PMESII elements and their complex network of relationships may be computationally realizable.

15. SUBJECT TERMS
PMESII, multiple representations, integrated reasoning, hybrid systems, social cognition, theory of mind

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Paul Bello
U	U	U	UL	26	19b. TELEPHONE NUMBER (*Include area code*)

Abstract

Assessing the effects of various actions on particular varieties of battlespace entities has been impeded by the absence of a representative computational model of the governing dynamics between so-called PMESII (Political, Military, Economic, Social, Information, and Infrastructure) elements in the theater of operations. Associated with each of these categories are formalisms suitable for computational implementation. Currently, we are lacking a way to describe the nonlinear dependence of each category on the others. We have proposed that the formal systems used to describe each of these categories can be described within an abstracted, yet common representational framework and therefore heterogeneous reasoning across the PMESII spectrum may be within our grasp. The Polyscheme architecture (Cassimatis 2002) was utilized to perform this integration, and to demonstrate that non-trivial reasoning about PMESII elements and their complex network of relationships may be computationally realizable. Further, emphasis was placed on the incorporation of a formal framework into Polyscheme within which to reason about the often ill-represented social and political dimensions of warfare. A proof-of-concept was developed demonstrating a novel capability to heterogeneously reason over information represented in the various PMESII battlespace partitions in support of linking commander's intent to quantifiable effects.

Table of Contents

List of Figures

Introduction

The Commander's Predictive Environment is a six-year joint effort between AFRL's Information Directorate, and its Human Effectiveness Directorate which is aimed at the development of a novel suite of tools to be used as advanced decision aides for decision-makers in our Air Operations Centers.

Some of the most difficult challenges in doing so involve a formalization of the direct connection between guidance from the commander (i.e. "commander's intent") in the form of a prioritized, temporally sequenced list of strategic goals to be accomplished to operational objectives and the tactical tasks used to accomplish them. All of this must be couched in terms of effects, and the actions used to generate them. Further complicating this issue is the fact that all effects aren't equal. Political and social effects for example, may only be able to be generated through the application of military or economic power at multiple levels of resolution (for example, taking actions against local governmental elements as opposed to taking action against an entire nation-state via say, economic sanctions). In order to make better ontological sense of this morass, we represent effects in six distinct domains: political, military, economic, social, infrastructure, and information-oriented (PMESII).

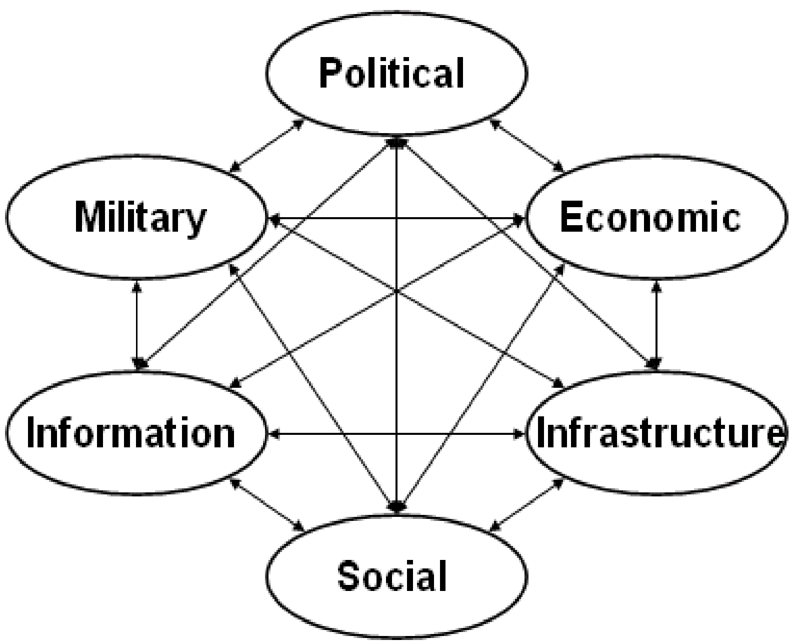

Figure 1 The PEMESII Web

We've also classified the types of action to be taken against an enemy system, when represented in this fashion. Actions are categorized as diplomatic, information-oriented, military, or

economic (DIME). Our challenge is to help discover the DIME actions, whether they be geopolitical or tactical which will generate strategic PMESII effects, and to do so in accordance with the multitude of constraints imposed by international, foreign, and domestic law.

Perhaps the most difficult challenge that lies before us is that of choosing a representation suitable for all the different varieties of data one would encounter across the PMESII spectrum. Individual researchers in this domain will have their proclivities and areas of specialization, but it seems rather far-fetched to assume that any one representational framework will be sufficient for representing things as dissimilar as say, economic pricing functions, and rules of engagement. Particular representational frameworks will have their associated strengths and weaknesses. For example, a neural network, or associated multi-variable regression may be able to succinctly represent nonlinear dependencies between economic indicators, while various rule-based formalizations would be much better suited to handle a set of rules for engagement. Merely representing data in this way, and hacking together an ad-hoc set of connections between individual data elements will not provide us a consistent framework for performing integration. To address this challenge we applied the Polyscheme architecture (Cassimatis 2002, 2004), which has been specifically developed to adequately handle problems of this sort. Polyscheme will be discussed in detail throughout the next section.

The other major research challenge which we began to address involves representing and reasoning about the political and social dimensions of warfare. Reasoning about this type of knowledge normally engenders certain semantic considerations, due to the fact that laws for example (which impose lots of constraints on the actions for certain agents) are breakable, and don't have a truth-value (true or false) per se. Rather, situations in which laws remain unbroken must be contrasted with situations in which some laws are broken. In fact, scenarios may abound where laws must be broken, due to extenuating circumstances, and the violations entailed by those scenarios must be effectively reasoned through. We will provide a knowledge representation format and an appropriate semantics for dealing with data of this variety. This research product was implemented in Polyscheme specialists and will tightly integrate with other specialists, allowing for legal/norm-based reasoning to be performed with all of the complications inherent in real-world reasoning (i.e. changing perceptions, and temporal constraints, among other things).

Our objectives

At the culmination of this effort, we wanted to have clearly demonstrated two distinct capabilities, supporting the overarching goal of providing automated assistance to the commander in the development of strategy, factoring in various PMESII considerations. The first of these is a demonstrated capability to reason about political and social phenomena, which are often represented as legislation and statements about norms. The development of an automated (or semi-automated human-in-the-loop) system for performing inferences over this type of knowledge will be the focus of this effort. The second capability is a modified version of

the Polyscheme architecture, providing a common substrate of basic operations within which multiple knowledge representation formats (neural networks, rule-based systems, and ontology, among others) may interoperate, without any discrepancy in semantics. This modified version of Polyscheme will implement our first research product (the automated reasoning engine for political/social knowledge), allowing us to draw inferences concerning both legal and social matters informed by data from other PMESII domains[1]. The application domain we've selected to put this new reasoning engine through its paces is reasoning about target nomination with respect to rules of engagement and a small ontology of desired effects. Since rules of engagement are readily available, there will be less overhead (in terms of salaried hours expended) involved, especially at such an early stage of the effort. A notional sketch of the final application associated with this project is presented in a subsequent section.

Methods, Assumptions, and Procedures

Since our fundamental challenge was integration, we first began by integrating reasoning about PMESII elements within Polyscheme. This required significant work into understanding epistemic cognition. We used a "Doctor's Without Borders" problem to test this work.

Polyscheme

Polyscheme was our vehicle for integration. We made significant modifications to the Polyscheme architecture during this work.

A Brief Overview of Polyscheme and Related Architectural Principles

Polyscheme (Cassimatis 2002) is a set of formal and computationally realizable principles which provide an explanation of how the human mind utilizes multiple representations during the execution of higher-order cognitive processes. Since many crucial aspects of Predictive Battlespace Awareness (PBA) and Intelligence Preparation of the Battlespace (IPB) rely on so-called "grey-matter (human-driven) fusion" of many different assessments, reports and other information, we believe that the same principles which motivate Polyscheme are at work in evaluating phenomena in the battlespace. The development of the DIME/PMESII descriptive framework lends further support for a need to develop decision-aides capable of simultaneously reasoning over multiple representations, as has been discussed previously in this document.

Two insights into the computational methods that underlie many computational cognitive

[1] For example, whether a particular group of agents should violate their obligation to be good citizens because they are enraged by the current economic situation, or have been incited by some information-bearing entity (say, the media or a powerful religious leader).

modeling paradigms enable an account of how these can be combined into an integrated theory of human cognitive architecture, and thus :

- Common function principle. Many reasoning and problem solving strategies can be composed of sequences of the same set of common functions.

- Multiple implementation principle. Each procedural unit can be executed using algorithms based on multiple representations.

Common function principle (CFP)

A survey of algorithms from many different subfields of computational cognitive science reveals that many of the same functions form the basis of many different kinds of algorithms. A first approximation of the list of these *common functions* includes:

- *Forward inference*. Given a set of beliefs, infer beliefs which follow from them.
- *Subgoaling*. Given the goal of establishing the truth of a proposition, P, make a subgoal of determining the truth values of propositions which would imply or falsify P.
- S*imulate alternate worlds*. Represent and make inferences about alternate, possible, hypothetical or counterfactual states of the world.
- *Identity matching*. Given a set of propositions about an object, find other objects which might be identical to it.

The CFP will be justified later by showing in detail how these common functions can implement a wide variety of algorithms. The following rough characterizations of two widely used algorithms in AI illustrates how completely different methodologies can be implemented using the same set of common functions (which are underlined). They will be more precisely characterized later.

- *Search*: “When uncertain about whether A is true, represent the world where A is true, perform forward inference, represent the world where A is not true, perform forward inference, if forward inference leads to further uncertainty, repeat.”
- *Stochastic simulation* (used widely in Bayes Network propagation): “When A is more likely than not-A, represent the world where A is true and perform forward inference in it more often than you do for the world where not-A is true.”

This illustrates how algorithms from different AI communities can be conceived of as different strategies for selecting the same basic set of operations.

Multiple implementation principle (MIP)

Each of the common functions can be implemented using multiple algorithms and representations. For example forward inference can conducted using at least these three mechanisms:

- *Production rule firing:* can involve matching a set of rules against a set of known facts to infer new facts.
- *Feed-forward neural networks:* take the facts represented by the activation of the input units, propagate these activations forward and output the facts represented by the values of the output units.
- *Memory*: The value of a slow-changing attribute at time t2 can be inferred by recalling its value at t1 when the interval between t1 and t2 is sufficiently brief.

This principle will help explain how lower-level processes make up and influence higher-order reasoning and problem solving. These insights motivate a computational framework that achieves a significant level of integration. High-order reasoning and problem solving algorithms can (thanks to the CFP) be integrated with each other by implementing each of them as sequences of common function executions. Executing these algorithms together in the same situation would simply require interleaving and or overlapping the sequences of common functions that make up each algorithm. Implementing common functions that make up higher-order reasoning and problem solving using multiple lower-order reasoning and problem-solving methods (thanks to the MIP) enables each step of lower-level computation to be integrated with and influence higher-level computation.

Architectural principles

We begin with the following hypothesis, enabled by the common function principle, and explore its consequences for mental architecture. Note that even though we present these hypotheses in terms of the mind, the analogy we draw is fairly simple: mountains of evidence exist for functional localization in the human brain. Different parts of our brain are primarily responsible for implementing various well-understood computational schemes (Granger 2005). The battlespace may be understood in the same way, being composed of different elements (political, military, economic, social, informational, and infrastructure), which interact in non-trivial ways, using their own distinct sets of representations.

Higher-order cognition-common function hypothesis

The mind implements higher-order reasoning and problem-solving strategies by executing sequences of common functions. The rest of this section describes how this hypothesis, together with what is empirically known about many cognitive processes, motivates a theory of human cognitive architecture that explains how the mind integrates many qualitatively different computational mechanisms. First, there are many reasons (reviewed by Baars, 1988) to believe that the mind has specialized processors, which are here called **specialists**, for perceiving, representing and making inferences about various aspects of the world. The MIP implies that each function can be implemented using multiple mechanisms. The next hypothesis suggests that the mind implements common functions with these specialists.

Specialist-common function implementation hypothesis

The mind is made up of specialized processors that implement the common functions, using computational mechanisms that are different from specialist to specialist. For example, the mind might have a place memory specialist that uses cognitive maps to keep track of the location of objects. It might also have a spatial relation specialist that uses a constraint mechanism to keep track of relations among mechanisms. The specialist common function hypothesis states that specialists such as these make forward inferences that constitute reasoning and problem solving, create subgoals, simulate alternate worlds and generate identity matches using their own specific computational mechanisms. This theory takes no definite position regarding whether the computation within these specialists is “encapsulated” from that within other specialists, though it does imply that in general specialists communicate through a central mechanisms about to be described. The CFP states that algorithms can be composed into sequences of common function executions. The MIP states that each function can be executed using multiple computational mechanisms. Assuming that the mind does execute higher-order inference algorithms using common functions, how many of its computational mechanisms (embodied in specialists) does it use at any moment? There are at least two reasons to believe the following answer to this question:

Integrative cognitive focus of attention hypothesis.

The mind uses all specialists simultaneously to execute each common function and the mind has an integrative cognitive focus of attention which at once forces the specialists to execute a particular common function on the current focus, integrates the results of this computation and distributes these results to each of the specialists. First, interference in the Stroop Effect (Stroop, 1935) between multiple kinds of processing (e.g., word and color recognition) suggests that multiple mental processes (i.e. specialists) engage sensory input simultaneously. That Stroop-like interference can be found with emotional, semantic and many other non-perceptual aspects of stimuli suggest that all, not just perceptual, specialists, simultaneously process the same information. Second, if interference in Stroop-like tasks is a result of the mind’s attempt to integrate information from multiple cognitive processes, then it is possible that that the mechanism for achieving this integration is a focus of attention. Treisman and Gelad (1980) have demonstrated that visual attention is the main medium for integrating information from multiple perceptual modalities. Polyscheme is based on the notion that just as the perceptual Stroop effect can be generalized to higher-level non-perceptual cognition, that integrative perceptual attention suggests the existence of a higher-level cognitive focus of attention that is the mind’s principle integrative mechanism. Whether the mind’s perceptual and higher-order focus of attention are the same is left for now as an open question. Since we are assuming that the mind implements higher-order reasoning and problem solving strategies by sequences of the individual functions specified in the CFP that are implemented by the computation that occurs during each focus of attention, the following hypothesis is implied:

Higher order cognition as attention selection hypothesis

The mind's mechanisms for choosing the cognitive focus of attention decide which higher-order reasoning and problem solving strategies it executes. As will be described below, these hypotheses collectively allow an explanation of how the mind integrates computational methods currently best modeled using very different computational mechanisms.

Polyscheme

Polyscheme embodies the theory of cognitive architecture outlined in the last section.

Formal preliminaries

Much of Polyscheme's operation involves communication among specialists that uses a simple, familiar propositional formalism. The actual details of the formalism itself are fairly arbitrary and not essential to the theory, though the existence of such a formalism is important. The formalism consists of propositions and truth values for those propositions. A proposition denotes that a predicate holds over an ordered set of objects during a temporal interval in a "world". R(x, y, t, w) states that the relation denoted by R holds over the entities denoted by x and y over the temporal interval denoted be t in the world
w. Worlds are entities that enable propositions about hypothetical worlds. For example, one can say that in the world, w, in which there is a hole in a cup, that the cup is now empty with the proposition, Empty(cup, now, w). Although propositions may seem cognitively implausible, it is possible to map propositions onto more the more familiar ACT-R chunk representation. For example, the most recent proposition could be reformulated as the following chunk: (c ISA empty-predicate object cup time now world w).

A specialist can indicate its belief or doubt in a proposition, called its *stance* on it, with a two-dimensional truth value. The first dimension is the degree of confidence it had in the proposition's truth and the second is the degree of confidence that it is false. The four confidence levels are C, L, l, m and indicate respectively that the truth or falsehood is certain, very likely, likely, or maybe the case. "?" indicates that there is no evidence either way. Thus, to say that a specialist takes the stance, (L,m), on a proposition, P, is to say that the specialist has evidence that P is very likely to be true and evidence that it might be false. The confidence levels do not correspond to numerical probabilities; only their relative order of confidence matters. This two-dimensional scheme enables Polyscheme to differentiate between the case where there is some evidence that X is true, say, (l,?), and the case where there is very good evidence that it is both true and false, e.g., (L,L). Both could be denoted with P(X) = .5 in a scalar probabilistic framework and would not capture the difference between the first situation of lukewarm confidence and the second, which indicates the potential of serious inconsistency. The truth value which represents the overall evidence all the specialists have about the truth of a proposition is called their *consensus* on that proposition.

Specialists

All specialists implement the following functions:

- *OpinonOn(Proposition)*. Specialists specify their stance on a Proposition.
- *ReportOpinions(Proposition,Opinions).* Specialists learn about the opinion of other specialists on a proposition.
- *PossibleMatches(Proposition).* For each entity in the proposition, specialists find entities that might be identical to it.
- *RequestAttractions*(). Specialists return a set of "attractions" (described later in this section) which specify propositions they would like to focus on.

Specialists implement the common functions:

- *Forward Inference*. When, from new information, a specialist is able to infer that a proposition, P, has truth value, TV, its RequestAttractions function will request attention for P. *OpinionOn(P)* will return the opinion that P has truth value, TV.
- *Subgoaling*. When a specialist can give information about the truth value of P if it knows the truth value of a set of propositions, S, it will include those in the set of propositions it requests through *RequestAttractions*.
- *Identity matching*. Identity hypotheses are generated by *PossibleMatches*.
- *Representing alternate worlds*. Since the propositions that are the input and output of these functions can regard alternate worlds, the specialists must be able to represent and make inferences about these possible worlds.

Focus of attention

Polyscheme embodies the integrative cognitive focus of attention hypothesis by including a focus of attention in the form of a proposition that all specialists focus on, store in memory and make inferences about simultaneously. At every time step, Polyscheme does the following:

1. Polyscheme's *focus manager* (described below) chooses a proposition to make the focus of attention.
2. Polyscheme collects the stances of the specialists on the proposition by calling their *OpinionOn* functions.
3. Polyscheme reports these stances to the other specialist using their *ReportOpinions* functions.
4. Using their own computational mechanisms, the specialists process this new information, make inferences and decide which propositions would help them make better information.
5. The focus manager collects the propositions specialists request to focus on by calling the specialists' *GetAttractions* functions.

Focus Manager

Polyscheme models the guiding of attention using two kinds of mechanisms. First, each specialist can indicate through its GetAttractions a set of *attractions* which each indicate a proposition it requests to focus on and a *charge function* that indicates for which times and how strongly this request is made. Second, Polyscheme has a *focus manager* which at each time step uses the attractions to focus on a proposition. Attractions can be conceived of as tuples (p,c), where p is a proposition to focus on and c is a charge function which indicates how strongly a proposition is to be focused on at a given time. For example, the charge function c = 1/et, indicates that the desire to focus on p is immediate and fades over time. The focus manager chooses the focus of attention at each time step using a two-step process:

1. It computes the charge of all the attractions.
2. It chooses the proposition, P, from the attraction with the highest charge.

How Polyscheme explains the integration of multiple cognitive processes

Having described how models based on different computational mechanisms, e.g. search, production rules and Bayesian networks, can be implemented in Polyscheme, it is now possible to explain how doing so explains how the mind integrates these mechanisms deeply in particular.

Polyscheme models the mind's execution of reasoning strategies as a sequence of foci and explains the integration of these strategies by the ability of sequences of foci from different strategies to be easily interleaved.

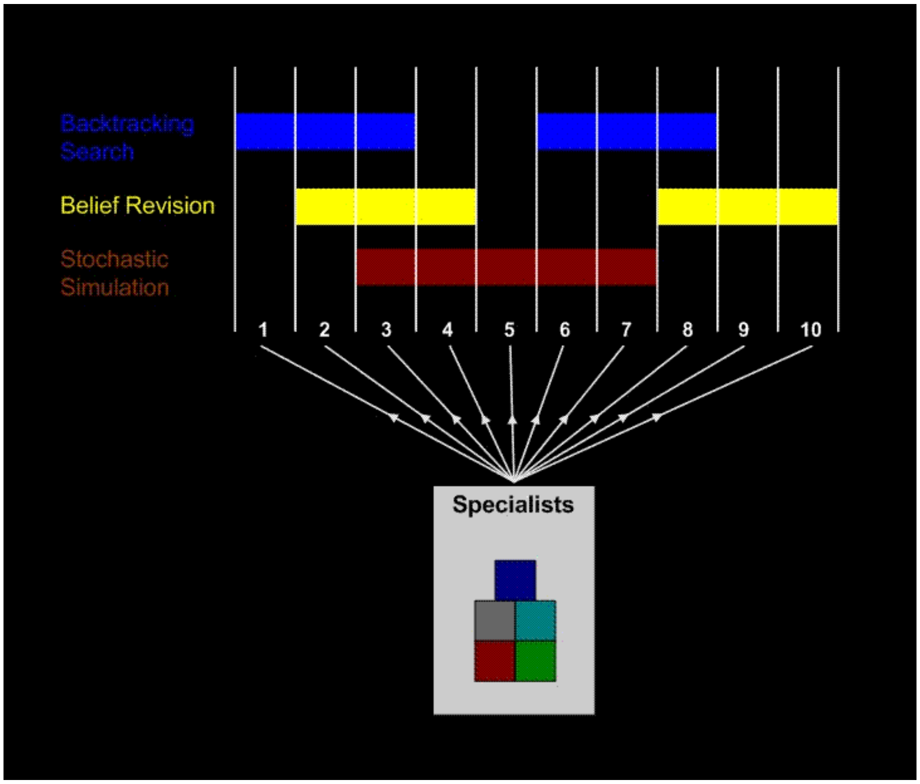

Figure 2. Interleaved Computation

For example, imagine a task requiring production rule firing and Bayesian network propagation and involves foci F1, F2… F11. Say production firing is implemented by focus F1, F3, F6, F7, F8, F9 and Bayesian network propagation through F2, F3, F4, F5, F6, F7, F10, and F11. Notice that the execution of both algorithms is *interleaved* so that inference made in the middle of, say, production rule matching can be used immediately in network propagation. Notice also that F3 and F6 are *shared* by both algorithms. This shows how Polyscheme enables models of *opportunism* by allowing computation involved in one strategy to be incorporated into another if the opportunity exists. Thus, Polyscheme provides an explanation of how the mind may flexibly combine reasoning and problem solving strategies by executing them as sequences of foci that can be interleaved and shared with each other.

If the mind implements reasoning and problem solving strategies using sequences of foci, all the computation is performed by the computation of the specialists during each focus. In other words, according to the Polyscheme theory, much higher-order cognition is nothing more then the guided focus of lower-level cognitive and perceptual processes. This helps explain how symbolic and serial cognitive processes are grounded (in the sense of (Harnad 1990)) in lower-level processes and to the extent these lower-level mechanism are sensorimotor, constitutes and embodied theory of higher-order reasoning. Also, since every focus of attention can be influenced by memory, perceptual and sensorimotor mechanisms, Polyscheme explains how reasoning can be interrupted or guided by these. Figure 3 illustrates the difference between

integration in Polyscheme and more strictly modular approaches to integration. On the left of a figure is a hypothetical cognitive model that includes models based on particular algorithms and data structures. Communication between these modules must be explicitly provided for. However, on the right, the execution of backtracking search as the focus of attention of each module simultaneously illustrates how Polyscheme explains how the mind integrates multiple cognitive processes in every step of reasoning continually and automatically.

Epistemic Cognition and the Cognitive Substrate

The essence of our work here was to show unequivocally that properly integrated domain-general mechanisms are sufficient for explaining how theory-of-mind emerges in normally developing children. As confirmation of its plausibility, our theory explains the developmentally synchronous successes on superficially different tasks that are normally used as benchmarks in assessing the normal development of theory-of-mind, and provides compelling explanations for why these capacities are limited in certain clinical populations, especially those stricken with autism.

While there are a number of different substrate elements that we believe are operative during theory of mind computations, three elements in particular seem to be critical for successes on these tasks: worlds/situations, similarity/identity, and causal rule usage. We will elaborate on each of these components with a focus on formally specifying their dynamics, and will justify their usage by appealing to relevant behavioral and neuroscientific findings. As we progress in our discussion and presentation of task models for false-belief, discrepant-belief, representational-change, appearance-reality, pretence usage, perspective-taking, and habituation studies in the theory-of-mind literature, we will use the specified dynamics to develop depictions of model execution traces. We have outlined a developmental trajectory of task-successes linked to the operation of substrate elements which comports well with a wide variety of empirical results in the ToM literature. A schematic representation of implicated substrate elements and the age-related task-successes which they enable is given below in Fig 3:

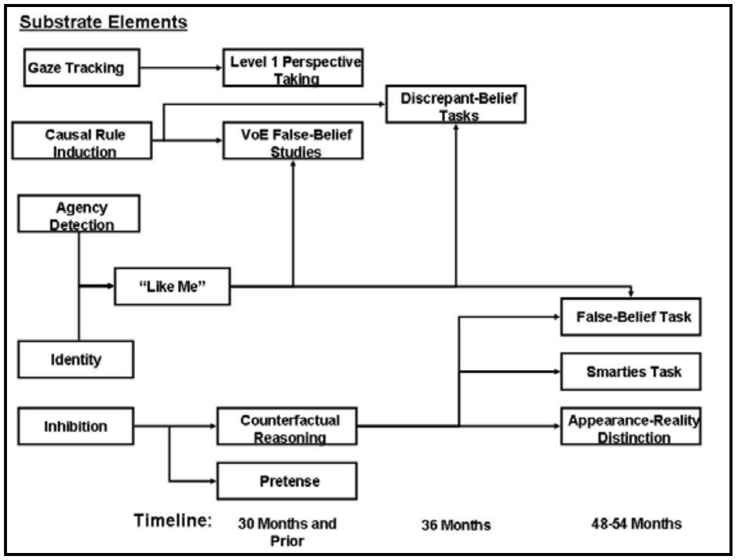

Figure 3. A Developmental Timeline of ToM Task Successes

As we will demonstrate in subsequent presentations of each task, most of the inference performed will be handled by physical reasoning mechanisms including search, forward inference, classification, and identity matching. However, in order to account for ToM development, we posit several new mechanisms and present justification for their inclusion in the substrate. We have added gaze-tracking, agency detection, and inhibitory processing to our basic set of substrate components. While not particularly crucial to this discussion, it can be noted that even agency-detection is really a species of categorization, and thus isn't technically an addition.

While none of these have been specifically required to perform the kinds of physical reasoning problems presented in (Cassimatis 2001), it seems reasonable to claim that they are part of our repertoire of cognitive mechanisms before our first year. Our ability to recognize the distinctness of both human faces, and the texture of human skin has been demonstrated in a number studies performed by Amanda Woodward and colleagues (Guajardo & Woodward 2004, Woodward 2005). Their experimental results demonstrate that judgments of intentional action are contingent in part upon judgments of agency – human agency in particular. Gaze following, while most likely not an innate faculty, emerges anywhere between 3 and 12 months of age in human neonates (Butterworth & Cochran 1980, Hood, Willen & Driver 1998). The underlying competencies driving the development of gaze-following are generally thought to include looking preferences (such as human faces), attention-related habituation strategies, simple reward-driven learning, and a social environment structured such that specific behaviors of other social creatures in the environment predicts the presence of interesting objects or events (Land, Mennie & Rusted 1999).

Causal Rules and Identification of Contingent Events

It is in this sort of structured environment involving feedings, playtime, and other turn-based activities that the induction of rules corresponding to appropriate environmental contingencies can successfully emerge. The rule "If *agent* has a head and body orientation consistent with the location of *object*, then *agent* sees *object*" emerges fairly straightaway, since it is part-and-parcel of almost every interaction that an infant would have with a caregiver. As the infant develops toward 18 months, direct attention to the perceptual apparatus of caregivers (primarily their eyes) is used as a reliable predictor of gaze direction and object reference (Butterworth & Jarrett, 1991). Theories regarding the ability of infants to infer the causal structure of their environment have been presented in (Gopnik et. al, 2004, Cohen et. al., 1998, Cohen & Amsel 1998). However, a debate continues over whether or not there exists an innate "causal module" which constantly polls the results of perception for contingency. Analogous to gaze-following phenomena, we choose to assume that such a capacity is already developed in the "youngest" version of our computational model.

In the models we will subsequently present, causal rules will be represented within a graphical formalism for ease of exposition (see Fig 4 for an example). Even with this being the case, it is useful to understand rules as they are actually implemented within the Polyscheme architecture.

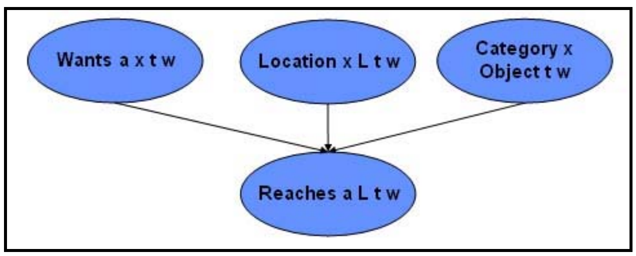

Figure 4: A Simple Causal Rule

Here is an example of a causal rule expressing the following: "if an agent *a* wants object *x* at some point in time *t* in world *w* and *x*'s location is at *L*, then *a* typically reaches for *x* at *L*." In Polyscheme's representational language, this rule would be expressed:

Wants(a, x, t, w) + Location(x, L, T, w) + Category(x, Object, t, w) ==> , Reaches(a, L, t, w).

Inference is performed by specialists over rules like this by matching the left-hand side against perceptions, and if all are fulfilled, detaching the right-hand side of the rule. This process can also work in reverse: if we know that it is not the case that *a* reaches at *L* at time *t* in world *w*, then at least one of the preconditions of reaching must be false. Either the agent doesn't want *x*, or *x* isn't an object, or *x* isn't located at *L*.

Self-Other Identity: Meltzoff's "Like-Me" Hypothesis

There is a fairly large body of evidence supporting the "Like Me" hypothesis put forth by Andrew Meltzoff and colleagues (Meltzoff 2005). Meltzoff's demonstration of imitation in

infants as young as 1 hour old has raised a host of interesting questions regarding potential mechanisms accounting for such behavior.

Alternate Worlds and Counterfactual Reasoning

We assume that children's minds are structured in such a way that they are capable of keeping track of how properties of objects in the world change through time. Consider the simple case of a ball rolling behind an occluder and emerging from the other side. Now consider a scenario in which a brick is dropped somewhere behind the occluder, and a subject is asked to predict whether or not the ball will emerge from the other side if rolled behind the occluder. A successful prediction in this scenario necessitates the consideration of at least two states of affairs, neither of which can be determined from observational evidence. The first of these is that the brick has fallen some distance behind the occluder which allows for the ball to pass between, and thus emerge on the right-hand side. The second is the scenario where the brick has fallen directly in line with the motion path of the ball, preventing it from emerging on the right.

A capacity for selective inhibition seems to be present in early infancy as well. This has been especially well-documented in the case of selective attention to salient visual cues (Amso & Johnson, 2004). Other examples of the gradual development of oculomotor inhibitory control can be found in (Scerif et. al., 2005).

DWB

The second technical objective was the identification of an appropriate scenario with highly sensitive initial conditions which allows for the demonstration of integrated reasoning across the PMESII spectrum and included epistemic inference. The scenario places a special emphasis on political and social phenomena which apparently are often troublesome to represent within traditional AI frameworks. In the first quarter, we identified a scenario which we deem to fit these criteria – and we have tentatively named it "doctors without borders." The scenario roughly approximates (and significantly embellishes) an actual operational incident in which a particular targeting decision to bomb a bridge had unexpected social consequences due to the fact that doctors without borders was using the river to provide medicine to upstream villages. The bombing hindered the timely delivery of medical aid to an area of post-war reconstruction, and thus altered public opinion of coalition intent. In the first quarter, we added economic, political and military ramifications to this scenario by couching it within the framework of a local governmental election between insurgent-backed and coalition-backed candidates.

In this quarter, we added an element of propaganda to our "doctors without borders" scenarios. Based on whether a population new of the knowledge and intent of the allies, their behavior attitudes would differ from scenario to scenario. We used the representational framework developed in our work on epistemic inference to model the influence of true and false beliefs on a populace's reactions to allied actions.

Results and Discussion

Polyscheme

This report began with the problem of combining models of cognitive processes based on difficult-to-integrate computational methods into a unified understanding of how the mind works. The computational function and multiple implementation principles, together with some empirical knowledge of how the mind works suggested a series of hypotheses about human cognitive architecture that together explain how the mind can integrate these different computational mechanisms. These principles have been embodied in a computational cognitive architecture, called Polyscheme. How accurately and comprehensively this theory explains human mental organization can only be determined by actually integrating cognitive models from various frameworks in Polyscheme and showing how this enables models of cognition in situations that are accurate and are difficult or impossible using individual frameworks alone. This paper's demonstration that models from many formerly difficult-to-integrates computational frameworks can be construed as an encouraging sign that this research program can achieve some success and insight.

Epistemic Cognition

The work resulted in a computationally realized model of social cognition to be used with the scenario to be identified. This model was developed in conjunction with Dr. Paul Bello (AFRL/IFSB). Data was collected and gathered across a wide range of child development experiments relating to "theory-of-mind" or the ability for humans (even young humans) to predict and explain the behavior of their conspecifics using mental-state terms such as beliefs, desires, and intentions. This data served as a set of constraints for the modification of the Polyscheme cognitive architecture in order to accommodate the modeling of these behaviors. Particularly, we focused on the notion of "false-beliefs" which seem to underlie both deceptive activity by adversaries, and distinctions that are drawn (or that fail to be drawn) by humans regarding the intentions of their conspecifics in relation to the view that they themselves take with respect to these intentions. As far as we are aware, the computational models developed for the first stage of this project represents a milestone in the cognitive architectures community – both in the scope of data that they are capable of accounting for, and in novelty, as their remains no extant literature on unified computational models of theory-of-mind development and usage. Moreover, these architectural modifications allow for a first-pass simulation of psychological operations within a PMESII framework.

Doctors without Borders

The result of our work on the Doctors Without Borders example was a system that could take into account political, military, economic, information and infrastructure considerations to make predictions about the effects of offensive actions. Because of the flexibility of computation enabled by the integration within the Polyscheme architecture, we were able to achieve a great deal of success in adapting to subtle change in the situation. For example, if our program considered a case in which the mosque was put on the same side of the river as the food and the city as opposed to the other side, then, even though from the point of view of computer

representation this was a very simple manipulation, the very profound real world consequence that would result was predicted by our system. These results demonstrate that Polyscheme, together with the epistemic cognition abilities we have added, are able to make a flexible predictions about the effects of offensive actions.

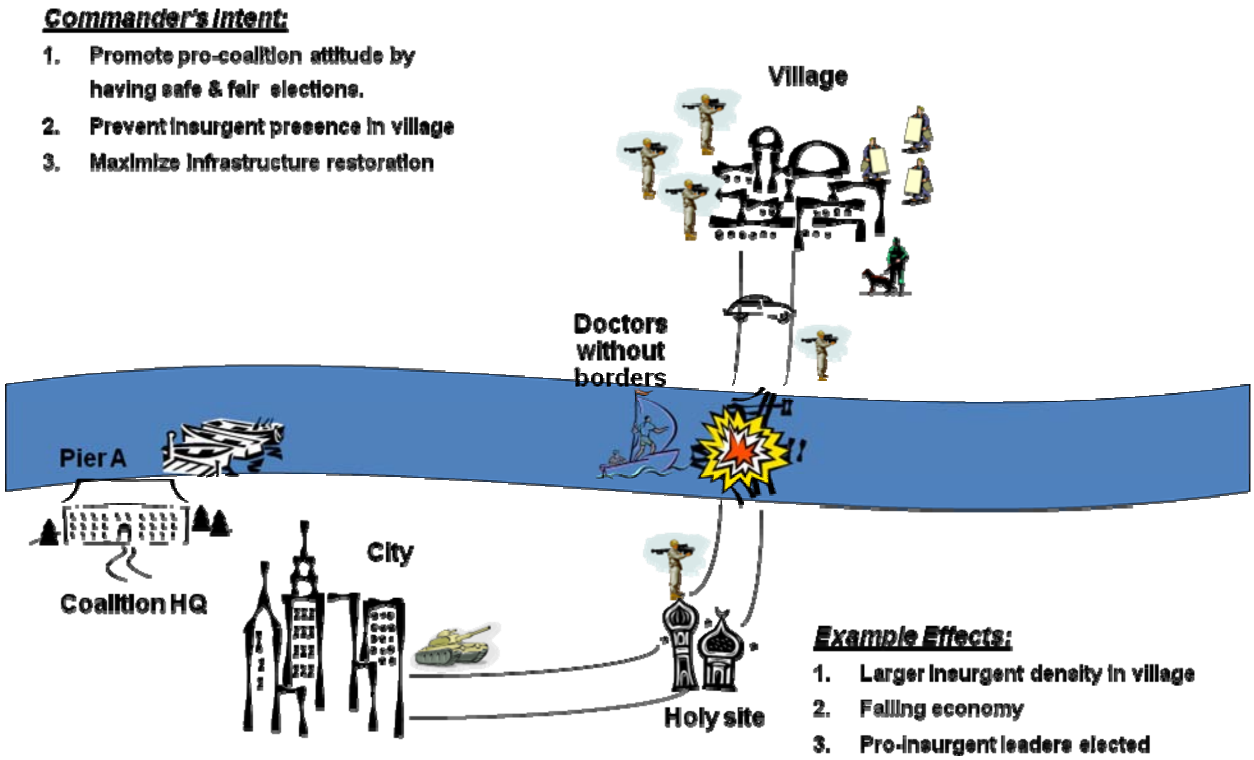

Figure 5. A Scenario Processed within Polyscheme

Integrated Development Environment for PEMSII modeling

One of the results of this work has been to create a graphical integrated development environment for the creation of PMESII models within Polyscheme. We have created an XML-based input format that allows users to easily add information in various knowledge representation formats into our PMESII models without any knowledge of how to program Polyscheme. These input files can be embedded inside a JBuilder or Eclipse source tree so that all the various programming tools – for example, code versioning systems and global search – can be applied to our PMESII model files .

Conclusions

Polyscheme as an integration medium

This report began with the problem of combining models of cognitive processes based on difficult-to-integrate computational methods into a unified understanding of how the mind works. The computational function and multiple implementation principles, together with some empirical knowledge of how the mind works suggested a series of hypotheses about human cognitive architecture that together explain how the mind can integrate these different computational mechanisms. These principles have been embodied in a computational cognitive architecture, called Polyscheme. How accurately and comprehensively this theory explains human mental organization can only be determined by actually integrating cognitive models from various frameworks in Polyscheme and showing how this enables models of cognition in situations that are accurate and are difficult or impossible using individual frameworks alone. This paper's demonstration that models from many formerly difficult-to-integrates computational frameworks can be construed as an encouraging sign that this research program can achieve some success and insight.

Conclusions from epistemic cognition work

Our work on other ToM tasks lends support to the generality of our claims here. We have used Polyscheme/ToM to model the appearance-reality distinction, the so-called "Smarties task," and the false belief task (Flavell 1986) using Polyscheme/ToM (Bello & Cassimatis, under review). Our model demonstrates how precise computational implementations can limit and help to resolve confusions in the interpretation of behavior in ToM tasks. Specifically, Polyscheme/ToM makes the following contributions:

- Explicit reasoning about beliefs is not necessary for success in false belief tasks. In our model, children can succeed at false belief tasks about, say, a cookie's location, without reifying (making explicit) the other person's belief about that cookie and associating it with the cookie in the world. In our model, children represent a state of affairs associated with Susan, but in that state of affairs the cookie is a cookie, not a representation or belief about a cookie. It is just a cookie in a different location. In terms of theory-theory, this means that these experiments do not require that a four year old's ontology includes beliefs. It merely requires that it includes states of affairs, objects in the world denoted as persons, and a way of relating them.
- Precision, in the form of computational models, can serve to clarify ambiguous experimental results. This is demonstrated by comparing our two models of action prediction in both the marker-finding (Wellman & Bartsch 1988) and cookie-finding scenarios (Wimmer & Perner 1983), and noting that the marker-finding task doesn't require the special functionality that alternate world inference offers, whereas the cookie-finding task does.
- More precise assessment of the magnitude of the developmental shift. Although the differences in our three-year old and four-year old models correspond to a significant shift in inferential power, they do so with surprisingly few changes in knowledge or mechanism. In

> addition to not requiring specific cognition about beliefs, desires and representing, the "theory" component in our model which explains performance on the false belief task is only one rule about how what one perceives affects his mental state and another rule about how such a mental state together with desire leads to action. The four-year-old innovation is to apply the alternate world mechanism (for which there is independent evidence (Dias & Harris, 1990) to representing other minds. No more sophisticated theoretical or conceptual apparatus is required.

ToM has been relatively unexplored in cognitive modeling. This work demonstrates that precise computable models can illuminate important issues in this literature. By precisely specifying what exactly constitutes children ToM, we have therefore been able to reduce the problem of explaining a broad shift in children's behavior between three and four years to the question of how they acquire/learn/develop one rule, modify a second rule, and begin to apply an alternate world mechanism they already possess toward representing other minds.

Conclusions from DWB.

Our work with the Doctors Without Border model using the advances within Polycheme and epistemic cognition demonstrate that the present multirepresentational approach is able to make a flexible predictions about the effects of offensive actions and therefore potentially increase commander's insight into the effects of ongoing and planned operations.

References

K. Arkoudas, S. Bringsjord, and P. Bello (2005). Toward Ethical Robots via Mechanized Deontic Logic. *In Proceedings of AAAI Fall 2005 Symposium on Machine Ethics*.

B. J. Baars (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University.

P. Bello (2005). *Toward a Logical Framework for Cognitive Effects-Based Operations*. Doctoral Dissertation, Rensselaer Polytechnic Institute, Troy, NY.

N. L. Cassimatis (2002). *Polyscheme: A Cognitive Architecture for Integrating Multiple Representation and Inference Schemes.* Doctoral Dissertation, Media Laboratory, Massachusetts Institute of Technology, Cambridge , MA .

N. L. Cassimatis (2004). Grammatical Processing Using the Mechanisms of Physical Inferences. In *Proceedings of the Twentieth-Sixth Annual Conference of the Cognitive Science Society*.

B. F. Chellas (1980). *Modal Logic: An Introduction*. Cambridge University Press, Cambridge, UK

R. Granger (2005). Engines of the Brain: The computational instruction set of human cognition. *Artificial Intelligence* (In press).

J. Horty (2001). *Agency and Deontic Logic*. Oxford University Press, Oxford, UK.

S. Harnad (1990). The Symbol Grounding Problem. *Physica* D 42: 335-346.

J. R. Stroop (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 622-643.

Y-H. Tan and L. W. N. van der Torre (1994). Multi-Preference Semantics for a Defeasible Deontic Logic. *In Proceedings of JURIX'94, Amsterdam, Holland*

A. M. Treisman and G. Gelade (1980). A feature integration theory of attention. *Cognitive Psychology, 12*, 97-136.

L. W. N. van der Torre (1997). *Reasoning about Obligations: Defeasibility in Preference-Based Deontic Logics*. Doctoral Dissertation, Erasmus University, Rotterdam, the Netherlands.

G.H. Von Wright (1980). Problems and Prospects in Deontic Logic. A Survey. In E. Agazzi (ed.), *Modern Logic,* Reidel Publishers, Dordrecht, The Netherlands, p. 399

Appendix – Publications resulting from this work

- N.L. Cassimatis (forthcoming). Adaptive Algorithmic Hybrids for Human-Level Artificial Intelligence. *Proceedings of the AGI Workshop 2006.* IOS Press. Eds. B. Goertzel and P. Wang.
- N.L. Cassimatis (2007). Reasoning as Cognitive Self-Regulation. In Integrated Models of Cognitive Systems. Oxford University Press. Ed. W. Gray.
- N. L. Cassimatis (2007). Some Computational Desiderata for Recognizing and Reasoning About the Intentions of Others. In *Proceedings of AAAI Spring Symposium*. Technical Report SS-07-03. pp. 1-6. Ed. George Ferguson.
- N.L. Cassimatis (2006). A Cognitive Substrate for Human-Level Intelligence. *AI Magazine.* Volume 27 Number 2. pp. 71-82.
- N.L. Cassimatis, E.K. Mueller, P.H. Winston (2006). Editors, Special Issue of AI Magazine on Achieving Human-Level Intelligence through Integrated Systems and Research. *AI Magazine*. Volume 27 Number 2.
- P. Bello, N.L. Cassimatis (2006). Developmental Accounts of Theory-of-Mind Acquisition: Achieving Clarity via Computational Cognitive Modeling. In *Proceedings of 28th Annual Conference of the Cognitive Science Society*.
- P. Bello & N.L. Cassimatis (2006). Understanding other Minds: A Cognitive Modeling Approach. In *Proceedings of the 7th International Conference on Cognitive Modeling*.
- N.L. Cassimatis (2006). Cognitive Science and Artificial Intelligence Have the Same Problem. Proceedings *of the 2006 AAAI Spring Symposium on Between a Rock and a Hard Place: Cognitive Science Principles Meet AI-Hard Problems.* Technical Report SS-06-02, eds. Christian Lebiere & Robert Wray, Eds. C. Lebiere & R. Wray.

Acronyms

CPE – Commander's predictive environment.
PMESII – Political, Military, Economic, Social, Infrastructure and Information
ToM – Theory of Mind